ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Complexity 21 (2005) 294–313

Journal of COMPLEXITY

www.elsevier.com/locate/jco

Approximation on anisotropic Besov classes with mixed norms by standard information☆

Gensun Fang[a], Fred J. Hickernell[b, c, *], Huan Li[a]

[a] *Department of Mathematics, Beijing Normal University, Beijing 100875, China*
[b] *Department of Applied Mathematics, Illinois Institute of Technology, 10 West 32nd Street, E1 Building, Room 208, Chicago, IL 60616-3793, USA*
[c] *Department of Mathematics, Peking University—Hong Kong Baptist University Joint Research Institute for Applied Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China*

Received 17 November 2004; accepted 5 January 2005
Available online 19 March 2005

Abstract

This article considers the approximation problem on periodic functions of anisotropic Besov classes with mixed norms using standard information. The asymptotic decay rates of the best algorithms in the worst-case setting are determined. An interpolating algorithm that attains this decay rate is given as well.

Keywords: Information-based complexity; Exact order; Interpolation; Optimal recovery

1. Introduction and main results

In the 1950s, under the influence of the work Kolmogorov [9], a new perspective on approximation theory developed. Classical approximation theory had already accumulated

☆ This work was partially supported by the Peking University—Hong Kong Baptist University Joint Research Institute for Applied Mathematics, by Hong Kong Research Grants Council Grant RGC/HKBU/2020/02P, by Hong Kong Baptist University Grant FRG/00-01/II-62, and by the National Natural Science Foundation of China Grant no. 10371009.

* Corresponding author. Department of Applied Mathematics, Illinois Institute of Technology, 10 West 32nd Street, E1 Building, Room 208, Chicago, IL 60616-3793, USA.
E-mail addresses: fanggs@bnu.edu.cn (G. Fang), fred@hkbu.edu.hk, hickernell@iit.edu (F.J. Hickernell)
URL: http://www.math.hkbu.edu.hk/∼fred.

0885-064X/$ - see front matter
doi:10.1016/j.jco.2005.01.001

a huge amount of results on the approximation by algebraic and trigonometric polynomials. Also, non-classical tools for approximation, such as splines, had began to permeate computational practice. Therefore, it was natural to ask how to compare different methods of approximation, how to determine an optimal method, and how to construct optimal or nearly-optimal algorithms in various settings. This led to the notions of widths, optimal recovery, and computational complexity. There have been many beautiful results on the exact asymptotic orders of these quantities for various function spaces, but there still remain many important open problems.

The extensive literature devoted to the widths, optimal recovery of functions, and computational complexity, includes the work of Heinrich [5], Micchelli and Rivlin [15,16], Novak [19], Pinkus [20], Ritter [21], Traub et al. [30] and Traub and Woźniakowski [31]. In particular, Heinrich [5] and Luo and Sun [12] have investigated the weak asymptotic order of the reconstruction of functions in classical Sobolev spaces using their values at n points. Kudryavtsev [10,11] has studied the same problem for non-periodic isotropic Besov spaces. His method to obtain upper bounds is different from the method used in this article.

It is well known that approximation plays a dominant role in the class of linear multivariate problems, and the results on approximation can be often used for other multivariate problems including integration. This article studies an approximation problem using standard information, that is the values of the function at selected points. Specifically, the function to be approximated lies in a multivariate anisotropic Besov space of periodic functions $B^{\mathbf{r}}_{\mathbf{p}\theta}$ defined below.

In this article $\mathbb{R}$ denotes the set of real numbers, and $\mathbb{R}_+$ the set of non-negative real numbers. Moreover, $\mathbb{Z}$ denotes the set of integers, $\mathbb{Z}_+$ the set of non-negative integers, and $\mathbb{N}$ the set of positive integers. The notation $\mathbb{R}^d$ denotes the space of d-dimensional vectors, and analogously for $\mathbb{R}^d_+$, etc. The periodic functions to be approximated are defined on the d-dimensional torus, $\mathbb{T}^d := [0, 2\pi]^d$.

Let $f(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, \dots, x_d) \in \mathbb{T}^d$, be a measurable, almost everywhere finite real function which is 2π-periodic in each variable. This function is in $L_{\mathbf{p}}(\mathbb{T}^d)$, for $\mathbf{p} = (p_1, p_2, \dots, p_d)$, $1 \leqslant p_j < \infty$, $j = 1, 2, \dots, d$, if

$$\|f\|_{\mathbf{p}} := \left(2\pi \int_0^{2\pi} \left(\frac{1}{2\pi} \int_0^{2\pi} \left(\frac{1}{2\pi} \int_0^{2\pi} |f(\mathbf{x})|^{p_1}\, dx_1\right)^{p_2/p_1} dx_2 \right.\right.$$
$$\left.\left. \cdots\right)^{p_d/p_{d-1}} dx_d\right)^{1/p_d} < \infty,$$

with the usual modification if some p_i are infinite, and the notation $L^{\mathbf{s}}_{\mathbf{p}}(\mathbb{T}^d)$ represents the space of all functions $f \in L_{\mathbf{p}}(\mathbb{T}^d)$ whose s_j partial derivatives $\partial^{s_j} f/\partial x_j^{s_j}$ on variate x_j, $j = 1, \dots, d$ are also in $L_{\mathbf{p}}(\mathbb{T}^d)$. The mixed norm [25, p. 21] is given by

$$\|f\|_{L^{\mathbf{s}}_{\mathbf{p}}} := \|f\|_{\mathbf{p}} + \sum_{j=1}^{d} \left\| \frac{\partial^{s_j} f}{\partial x_j^{s_j}} \right\|_{\mathbf{p}},$$

where $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{Z}_+^d$. In this norm, $L_{\mathbf{p}}^{\mathbf{s}}(\mathbb{T}^d)$ is a Banach space, and $L_{\mathbf{p}}(\mathbb{T}^d) := L_{\mathbf{p}}^{\mathbf{0}}(\mathbb{T}^d)$. In this article the error between the original function and its approximation is measured in the space $L_{\mathbf{q}}^{\mathbf{s}}(\mathbb{T}^d)$.

The Besov space of periodic functions is defined in terms of difference quotients. For any $k_i \in \mathbb{N}$ the k_ith difference of the function $f(\mathbf{x})$ at the point $\mathbf{x}$ for x_i with step h_i is denoted by

$$\Delta_{h_i}^{k_i} f(\mathbf{x}) = \sum_{j=0}^{k_i} (-1)^{k_i+j} \binom{k_i}{j} f(x_1, \dots, x_{i-1}, x_i + jh_i, x_{i+1}, \dots, x_d).$$

This difference can be used to define a modulus of continuity of a function f in $L_{\mathbf{p}}(\mathbb{T}^d)$, namely

$$\omega_{k_i}(f, t_i)_{\mathbf{p}} := \sup_{|h_i| \leqslant t_i} \|\Delta_{h_i}^{k_i} f\|_{\mathbf{p}}, \qquad t_i > 0,\ i = 1, \dots, d.$$

In this article when comparing vectors, e.g., $\mathbf{p}$, $\mathbf{q} \in \mathbb{R}^d$, the notation $\mathbf{p} < \mathbf{q}$ means $p_i < q_i,\ i = 1, \dots, d$, and the notation $\mathbf{p} \leqslant \mathbf{q}$ means $p_i \leqslant q_i,\ i = 1, \dots, d$. Moreover, $\mathbf{1}$ denotes a vector of ones, and $\boldsymbol{\infty}$ is the vector whose elements are infinite. The inequality $\mathbf{p} < \boldsymbol{\infty}$ means that all the elements of $\mathbf{p}$ are finite. The following definition describes the anisotropic Besov space of periodic functions.

Definition 1 (*Nikolskii [18, pp. 161–163, p. 217]*). Let $\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{N}^d$ and $\mathbf{r} = (r_1, \dots, r_d) \in \mathbb{R}_+^d$, with $\mathbf{k} > \mathbf{r}$, $1 \leqslant \theta \leqslant \infty$, and $\mathbf{1} \leqslant \mathbf{p} \leqslant \boldsymbol{\infty}$. A function f is said to lie in the anisotropic Besov space $B_{\mathbf{p}\theta}^{\mathbf{r}}(\mathbb{T}^d)$ if and only if $f \in L_{\mathbf{p}}(\mathbb{T}^d)$ and

$$\|f\|_{b_{x_i \mathbf{p}\theta}^{r_i}(\mathbb{T}^d)} = \begin{cases} \left(\int_0^\infty \left(\frac{\omega_{k_i}(f, t_i)_{\mathbf{p}}}{t_i^{r_i}} \right)^\theta \frac{dt_i}{t_i} \right)^{1/\theta} < \infty, & 1 \leqslant \theta < \infty; \\ \sup_{t_i > 0} \frac{\omega_{k_i}(f, t_i)_{\mathbf{p}}}{t_i^{r_i}} < \infty, & \theta = \infty, \end{cases}$$

for $i = 1, \dots, d$. The linear space $B_{\mathbf{p}\theta}^{\mathbf{r}}(\mathbb{T}^d)$ is a Banach space with the norm

$$\|f\|_{B_{\mathbf{p}\theta}^{\mathbf{r}}(\mathbb{T}^d)} := \|f\|_{\mathbf{p}} + \sum_{i=1}^d \|f\|_{b_{x_i \mathbf{p}\theta}^{r_i}(\mathbb{T}^d)}$$

and is called an anisotropic Besov space. For any Banach space X the unit ball centered at the origin is denoted $\mathcal{B}(X)$, and is defined as $\{f \in X : \|f\|_X \leqslant 1\}$. Specifically, the unit ball of this Besov space centered at the origin is defined as

$$S_{\mathbf{p}\theta}^{\mathbf{r}}(\mathbb{T}^d) := \mathcal{B}(B_{\mathbf{p}\theta}^{\mathbf{r}}(\mathbb{T}^d)) := \{f \in L_{\mathbf{p}}(\mathbb{T}^d) : \|f\|_{B_{\mathbf{p}\theta}^{\mathbf{r}}(\mathbb{T}^d)} \leqslant 1\}.$$

The ambiguity in choice of $\mathbf{k}$ with $\mathbf{k} > \mathbf{r}$ is unimportant, since different $\mathbf{k}$ correspond to equivalent semi-norms. It follows from [18, p. 153] that $B_{\mathbf{p}\infty}^{\mathbf{r}}(\mathbb{T}^d)$ coincides with the classical Hölder–Nikolskii space $H_{\mathbf{p}}^{\mathbf{r}}(\mathbb{T}^d)$. If $r_1 = \dots = r_d = r$, these spaces of functions are called isotropic. The anisotropic Besov space with mixed norms is very important in numerical analysis, optimal recovery, differential and integral equations, probability and mathematical physics (e.g., see [1,3,18,25,29]). In particular, the bilinear approximation problem on classes of functions with mixed norms is closely related to the calculation of the widths of these classes [26, Chapter 4,14]. By using Besov spaces instead of classical Sobolev spaces Dahlke and DeVore [3] were able to improve the regularity assertions for solutions to boundary value problems for the Laplace operator on a Lipschitz domain. Moreover, there are various situations in partial differential equations where one can control the solution $u = u(x, t)$ with respect to a norm like

$$\int_0^T \left(\int^2_{\Omega|u(x,t)|} dx \right)^{1/2} dt.$$

Denote by $C(\mathbb{T}^d)$ the space of real, continuous, and periodic functions on $\mathbb{T}^d$. Let X be a normed linear space of real functions defined on $\mathbb{T}^d$, and $K \subset C(\mathbb{T}^d) \cap X$. The quantity

$$d(K) := \sup_{f,g\in K} \|f - g\|_X$$

is called the diameter of K.

Function approximation is usually based on standard information, i.e., the values of the function at some points. Mathematically, this may be written as $I_{B_n}(f) := \{f(\mathbf{b}_i)\}_{i=1}^n$, where I is called a sampling operator, the design is $B_n = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\} \subset \mathbb{T}^d$, and $n \in \mathbb{N}$. The net width [1] or the nth minimum information diameter [31, p. 48] is defined as

$$\Delta_n(K, X) := \inf_{B_n} \sup_{f\in K} d\left(\left(I^{-1}_{B_n}(I_{B_n}(f)) \right) \cap K \right), \tag{1}$$

where $I^{-1}_{B_n}(I_{B_n}(f))$ denotes the set of all functions that share the same values as f on the design.

A mapping $\phi : \ I_{B_n}(K) \to X$ is called an algorithm, and $\phi(I_{B_n} f)$ is the approximation of f in X. Denote by Φ_{B_n} the set of all algorithms using n pieces of standard information on K. The quantity

$$e_n(K, X) := \inf_{B_n} \inf_{\phi\in\Phi_{B_n}} \sup_{f\in K} \|f - \phi(I_{B_n} f)\|_X \tag{2}$$

is called the nth minimum intrinsic error of the optimal recovery of the set K in the space X. Denote by $\Phi^{\mathrm{L}}_{B_n}$ the set of all linear algorithms ϕ. By analogy one may define $e^{\mathrm{L}}_n(K, X)$, the nth minimum linear intrinsic error by substituting $\Phi^{\mathrm{L}}_{\mathbf{b}_n}$ in the place of Φ_{B_n} in the right-hand side of (2). If K is a center symmetrical and convex subset of X, then by Traub et al. [30, p. 67] it follows that

$$\tfrac{1}{2}\, \Delta_n(K, X) \leqslant e_n(K, X) \leqslant e^{\mathrm{L}}_n(K, X).$$

Usually K is chosen to be $\mathcal{B}(Y)$, the unit ball in some Banach space Y centered at the origin.

Our main results are for the case where $K = S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d) := \mathcal{B}(B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d))$, the unit ball in the Besov space, and $X = L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)$. The convergence rates of the approximation errors depend essentially on

$$g(\mathbf{r}) := \left(\sum_{i=1}^{d} \frac{1}{r_i}\right)^{-1}, \tag{3}$$

where $dg(\mathbf{r})$ is the harmonic average of the smoothness parameters $r_1, \ldots, r_d$. For the case $r_1 = \cdots = r_d = r$ it follows that $g(\mathbf{r}) = r/d$. In this article, it is always assumed that

$$g(\mathbf{r}) > \max_{i=1,\ldots,d} 1/p_i.$$

This ensures that $B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)$ is embedded into $C(\mathbb{T}^d)$ by a Sobolev-type embedding theorem [29, Chapter 2, Section 3].

In proving convergence rates it is convenient to use the notations $\ll$ and $\asymp$. For two nonnegative sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$. The relation $a_n \ll b_n$ means that there is a positive number C such that $a_n \leqslant C b_n$ for all n. The weak equivalence relation $\asymp$ means that $a_n \ll b_n$ and $b_n \ll a_n$.

The main results of this article are as follows:

Theorem 2. *Consider any vectors* $\mathbf{k} = (k_1, k_2, \ldots, k_d) \in \mathbb{N}^d$, $\mathbf{r} = (r_1, r_2, \ldots, r_d) \in \mathbb{R}^d_+$, $\mathbf{s} = (s_1, s_2, \ldots, s_d) \in \mathbb{Z}^d_+$, *with* $\mathbf{s} < \mathbf{r} < \mathbf{k}$ *and* $\mathbf{1} < \mathbf{r}$. *Moreover, let* $1 \leqslant \theta \leqslant \infty$ *and* $n \in \mathbb{N}$. *If either* $\mathbf{1} < \mathbf{q} \leqslant \mathbf{p} < \infty$ *or* $\mathbf{1} = \mathbf{q} < \mathbf{p} < \infty$, *then it follows that*

$$\Delta_n\left(S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d), L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)\right) \asymp e_n\left(S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d), L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)\right) \asymp e^L_n\left(S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)), L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)\right) \asymp n^{-g(\mathbf{r})(1-\beta)}$$

where

$$\beta = \max_{i=1,\ldots,d} s_i/r_i. \tag{4}$$

The interpolation algorithm $\mathcal{D}_{\mathbf{N}}(f)$ *defined in* (5) *in the next section is an optimal algorithm in the sense of exact order.*

Theorem 3. *Make the same assumptions on* $\mathbf{k}, \mathbf{r}, \mathbf{s}, \theta$, *and* n *as in Theorem* 2. *If* $\sum_{i=1}^{d} (1/p_i - 1/q_i)1/r_i < 1$, *and* $\mathbf{1} < \mathbf{p} \leqslant \mathbf{q} < \infty$ *or* $\mathbf{1} = \mathbf{p} < \mathbf{q} < \infty$, *then*

$$\Delta_n\left(S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)), L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)\right) \asymp e_n\left(S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d), L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)\right) \asymp e^L_n\left(S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)), L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)\right) \asymp n^{-g(\mathbf{r})(1-\beta)+g(\mathbf{r})\sum_{i=1}^{d}(1/p_i-1/q_i)1/r_i}$$

and the interpolation algorithm $\mathcal{D}_{\mathbf{N}}(f)$ *defined in* (5) *is an optimal algorithm in the sense of exact order.*

2. Interpolation via the Dirichlet kernel

Denote by

$$D_n(x) = \sum_{|k| \leqslant n} e^{\imath kx} = \frac{\sin((n+1/2)x)}{\sin(x/2)},$$

the classical Dirichlet kernel in one dimension, where $\imath = \sqrt{-1}$. The d-dimensional analog of this kernel is defined by

$$D_{\mathbf{N}}(\mathbf{x}) = \prod_{j=1}^{d} D_{N_j}(x_j),$$

where $\mathbf{N} = (N_1, \ldots, N_d) \in \mathbb{Z}_+^d$.

This Dirichlet kernel may be used to define an interpolation algorithm. Define the index set

$$P(\mathbf{N}) = \left\{ \mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}_+^d : \mathbf{n} \leqslant 2\mathbf{N} \right\}$$

and the points that comprise the design as

$$\boldsymbol{\xi}_{\mathbf{n}} := \left(\frac{2\pi n_1}{2N_1+1}, \ldots, \frac{2\pi n_d}{2N_d+1} \right) \in \mathbb{T}^d, \quad \mathbf{n} \in P(\mathbf{N}).$$

The Dirichlet approximation operator is defined as

$$\mathcal{D}_{\mathbf{N}}(f)(\mathbf{x}) := \prod_{j=1}^{d} (2N_j+1)^{-1} \sum_{\mathbf{n} \in \mathbf{P}(\mathbf{N})} f(\boldsymbol{\xi}_{\mathbf{n}}) D_{\mathbf{N}}(\mathbf{x} - \boldsymbol{\xi}_{\mathbf{n}}). \tag{5}$$

This approximation operator uses $n = (2N_1+1)\cdots(2N_d+1)$ function values.

The Dirichlet approximation operator has a couple of important properties [23]. First, it interpolates functions defined on $\mathbb{T}^d$, on the design $\{\boldsymbol{\xi}_{\mathbf{n}}\}$, i.e.,

$$f(\boldsymbol{\xi}_{\mathbf{n}}) = \mathcal{D}_{\mathbf{N}}(f)(\boldsymbol{\xi}_{\mathbf{n}}) \quad \text{for all } \mathbf{n} \in P(\mathbf{N}).$$

Second, if $T(\mathbf{N}, d)$ is the set of all multivariate trigonometric polynomials with wave numbers $\mathbf{k}$ satisfying $-\mathbf{N} \leqslant \mathbf{k} \leqslant \mathbf{N}$, then $\mathcal{D}_{\mathbf{N}}(t) = t$ for all $t \in T(\mathbf{N}, d)$.

In one dimension ($d = 1$) the interpolation operator $\mathcal{D}_n(f)$ is well studied. For example, the classical periodic Sobolev class $W_p^r(\mathbb{T})$ of periodic functions f permit the representation

$$f(x) = \frac{1}{2\pi} \int_0^{2\pi} \varphi(t) F_r(x-t, r)\, dt, \quad \|\varphi\|_p \leqslant 1,$$

where

$$F_r(x, \alpha) = 1 + 2 \sum_{k=1}^{\infty} k^{-r} \cos\left(kx - \frac{\alpha\pi}{2}\right), \quad \alpha \in \mathbb{R},$$

is the Bernoulli kernel. It is well known that (see the survey of Nikolskii [17])

$$\sup_{f\in\mathcal{B}(W^r_\infty(\mathbb{T}))}\|f-\mathcal{D}_n(f)\|_\infty \asymp n^{-r}\ln n,\quad n=2,3,\ldots.$$

As was noted by Temlyakov [27] (see also [28]) the results of Hristov [7] and Ivanov [8] imply that for $1/r<p$ and $1<p<\infty$,

$$\sup_{f\in\mathcal{B}(W^r_p(\mathbb{T}))}\|f-\mathcal{D}_n(f)\|_p \asymp n^{-r},\quad n=1,2,\ldots.$$

In addition to the approximation operator based on the Dirichlet kernel, there is an analogous operator,

$$\mathcal{V}_n(f):=\frac{1}{4n}\sum_{j=1}^{4n} f(\zeta_j)V_n(x-\zeta_j),\quad \zeta_j=\frac{\pi j}{2n},$$

based on the Vallee–Poussin kernel,

$$V_n(x):=1+2\sum_{j=1}^{n}\cos jx+2\sum_{j=n+1}^{2n}\frac{2n-j}{n}\cos jx,$$

that has been also studied by many authors. Among others, Temlyakov [27,29] proved that

$$e_n(\mathcal{B}(W^r_p(\mathbb{T})),L_q(\mathbb{T}))\asymp \sup_{f\in\mathcal{B}(W^r_p(\mathbb{T}))}\|f-\mathcal{V}_n(f)\|_q\asymp n^{-r+(1/p-1/q)_+}$$

for all $1\leqslant p,q\leqslant\infty$ and $r>1/p$, where $x_+=\max\{0,x\}$.

Consider now the multidimensional case. It follows from [29, Chapter 2, Section 3] that the periodic, anisotropic Sobolev space $W^{\mathbf{r}}_{p,\boldsymbol{\alpha}}(\mathbb{T}^d)$, $\mathbf{r}=(r_1,\ldots,r_d)>\mathbf{0}$, $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_d)\in\mathbb{R}^d$ consists of functions f which have the following integral representation for each $1\leqslant j\leqslant d$,

$$f(\mathbf{x})=\frac{1}{2\pi}\int_0^{2\pi}\phi_j(x_1,\ldots,x_{j-1},y,x_{j+1},\ldots,x_d)F_{r_j}(x_j-y,\alpha_j)\,dy,$$
$$\|\phi_j\|_p\leqslant 1$$

and the periodic Nikolskii space $NH^{\mathbf{r}}_p(\mathbb{T}^d)$ is the set of functions $f\in L_p(\mathbb{T}^d)$ such that for each $l_j=\lfloor r_j\rfloor+1$, $j=1,\ldots,d$ the following relations hold:

$$\|f\|_p\leqslant 1,\quad \|\Delta_h^{l_j}(f)\|_p\leqslant |h|^{r_j},$$

where $\Delta_h^{l_j}$ is the l_jth difference with step h in the variable x_j, $j=1,\ldots,d$. As usual this space is isotropic in the case $r_1=\cdots=r_d=r$. Temlyakov [28,29] obtained the following results:

$$e_n(\mathcal{B}(W^{\mathbf{r}}_{p,\boldsymbol{\alpha}}(\mathbb{T}^d)),L_q(\mathbb{T}^d))$$

$$\asymp \sup_{f\in\mathcal{B}(W^{\mathbf{r}}_{p,\boldsymbol{\alpha}}(\mathbb{T}^d))} \|\mathcal{V}_{\mathbf{n}}(f)-f\|_q \asymp e_n(\mathcal{B}(NH^{\mathbf{r}}_p(\mathbb{T}^d)), L_q(\mathbb{T}^d))$$
$$\asymp \sup_{f\in\mathcal{B}(NH^{\mathbf{r}}_p(\mathbb{T}^d))} \|\mathcal{V}_{\mathbf{n}}(f)-f\|_q \asymp n^{-g(\mathbf{r})+(1/p-1/q)_+}, \tag{6}$$

where $1\leqslant p, q\leqslant\infty$, $g(\mathbf{r}) > 1/p$, and $\mathcal{V}_{\mathbf{n}}(f)$ with $\mathbf{n}\in\mathbb{N}^d$ is defined in a product fashion as was done for the Dirichlet approximation operator.

Taking $r_1=\cdots=r_d=r$ gives the convergence rate for the classical (isotropic) Sobolev space:

$$e_n(\mathcal{B}(W^r_p(\mathbb{T}^d)), L_q(\mathbb{T}^d)) \asymp \sup_{f\in\mathcal{B}(W^r_p(\mathbb{T}^d))} \|\mathcal{V}_{\mathbf{n}}(f)-f\|_q \asymp n^{-r/d+(1/p-1/q)_+} \tag{7}$$

(see [5,2,6] for more details). On the other hand, Temlyakov [27,28, Chapter 3] also obtained some results on recovering functions by standard information for the Sobolev space $MW^{\mathbf{r}}_{p,\alpha}(\mathbb{T}^d)$ with bounded mixed derivative which is defined by

$$f(\mathbf{x}) = (2\pi)^{-d}\int_{\mathbb{T}^d}\varphi(\mathbf{y})\prod_{i=1}^d F_{r_j}(x_j-y_j,\alpha_j)\,dy_1\dots dy_d, \quad \|\varphi\|_p\leqslant 1.$$

Remark 4. Temlyakov's results in (6) using the Dirichlet operator follow from Theorems 2 and 3 by taking $\theta=\infty$, $\mathbf{s}=(0,\dots,0)$, $p_1=\cdots=p_d=p$, $q_1=\cdots=q_d=q$. The classical results in (7) follow by also taking $r_1=\cdots=r_d=r$.

3. Upper bounds

Theorems 2 and 3 give both upper and lower bounds on the asymptotic decay rates of the approximation errors. This section constructs upper bounds by determining the asymptotic decay rate of approximation using the Dirichlet interpolation algorithm (5). The proof is given in a series of lemmas.

The following lemma gives an estimates of the mixed norm of a trigonometric polynomial by its mixed lattice norm, which is a Marcinkiewicz-type inequality, and plays an important role in the proof of Theorem 2. Some additional notation is needed to state it concisely. Let

$$l(\mathbf{N},d) = \{\mathbf{a}:\mathbf{a}=\{a_{\mathbf{n}}\}, \mathbf{n}=(n_1,n_2,\dots,n_d),\ \mathbf{1}\leqslant\mathbf{n}\leqslant\mathbf{N}\}.$$

For $\mathbf{a}\in l(\mathbf{N},d)$, its mixed norm is defined as following:

$$\|\mathbf{a}\|_{\mathbf{p},\mathbf{N}} =: \left(\frac{1}{N_d}\sum_{n_d=1}^{N_d}\left(\cdots\left(\frac{1}{N_1}\sum_{n_1=1}^{N_1}|a_{\mathbf{n}}|^{p_1}\right)^{p_2/p_1}\cdots\right)^{p_d/p_{d-1}}\right)^{1/p_d}$$
$$= \|\mathbf{a}\|_{\mathbf{p}}\prod_{j=1}^d N_j^{-1/p_j}, \tag{8}$$

where

$$\|\mathbf{a}\|_{\mathbf{p}} = \left(\sum_{n_d=1}^{N_d} \left(\cdots \left(\sum_{n_1=1}^{N_1} |a_{\mathbf{n}}|^{p_1} \right)^{p_2/p_1} \cdots \right)^{p_d/p_{d-1}} \right)^{1/p_d} < \infty.$$

Lemma 5 (*Marcinkiewicz inequality, Schmeisser [23]*). *Let $t \in T(\mathbf{N}, d)$. Then for any $\mathbf{1} < \mathbf{p} < \infty$,*

$$\|t\|_{\mathbf{p}} \leqslant C_{d,\mathbf{p}} \|\{t(\xi_{\mathbf{n}})\}_{\mathbf{n} \in \mathbf{P}(\mathbf{N})}\|_{\mathbf{p},2\mathbf{N}},$$

where $C_{d,\mathbf{p}}$ is a number depending only on d, and $\mathbf{p}$.

The proof of this lemma is given in [23], and in a more general form in [22].

Lemma 6. *Suppose that $f \in L_{\mathbf{p}}^{\mathbf{r}}(\mathbb{T}^d)$, $\mathbf{r} \geqslant \mathbf{1}$, and $\mathbf{1} \leqslant \mathbf{p} < \infty$. Then it follows that*

$$\begin{aligned} &\left(\sum_{k_d=1}^{2n_d} \frac{2\pi}{n_d} \left(\cdots \left(\sum_{k_1=1}^{2n_1} \frac{2\pi}{n_1} |f(2k_1\pi/n_1, \ldots, 2k_d\pi/n_d)|^{p_1} \right)^{p_2/p_1} \cdots \right)^{p_d/p_{d-1}} \right)^{1/p_d} \\ &\quad \leqslant \|f\|_{\mathbf{p}} + \sum_{1 \leqslant i \leqslant d} \frac{2\pi}{n_i} \left\| \frac{\partial f}{\partial x_i} \right\|_{\mathbf{p}} + \sum_{1 \leqslant i \leqslant j \leqslant d} \frac{2\pi}{n_i} \frac{2\pi}{n_j} \left\| \frac{\partial^2 f}{\partial x_i x_j} \right\|_{\mathbf{p}} + \cdots \\ &\qquad + \prod_{i=1}^{d} \frac{2\pi}{n_i} \left\| \frac{\partial^d f}{\partial x_1 \cdots \partial x_d} \right\|_{\mathbf{p}}. \end{aligned}$$

Proof. For the dimension $d = 1$, the lemma is proved in [4,24], the case $d > 1$ with mixed norm, can be proved in an analogous way. But for the sake of readability, we give the proof in some detail.

For ease of notation, let $v_{i,k_i} = 2\pi k_i/n_i$. For $d = 1$ the mean value theorem implies the existence of an $\eta_{k_1} \in (v_{1,k_1-1}, v_{1,k_1})$, such that

$$|f(\eta_{k_1})| = \frac{n_1}{2\pi} \int_{v_{1,k_1-1}}^{v_{1,k_1}} |f(x_1)| \, dx_1.$$

This leads to an upper bound for $f(v_{i,k_1})$ in terms of integrals of f and its first derivative with respect to x_i:

$$\begin{aligned} |f(v_{1,k_1})| &\leqslant |f(\eta_{k_1})| + |f(v_{1,k_1}) - f(\eta_{k_1})| \\ &\leqslant \frac{n_1}{2\pi} \int_{v_{1,k_1-1}}^{v_{1,k_1}} \left\{ |f(x_1)| + \frac{2\pi}{n_1} \left| f'(x_1) \right| \right\} dx_1. \end{aligned} \tag{9}$$

Continuing this argument iteratively for higher dimensions yields

$$\begin{aligned}
&|f(v_{1,k_1},\ldots,v_{d,k_d})| \\
&\quad\leqslant \frac{n_1\cdots n_d}{(2\pi)^d}\int_{v_{1,k_1-1}}^{v_{1,k_1}}\cdots\int_{v_{d,k_d-1}}^{v_{d,k_d}}\Bigg\{|f|+\sum_{i=1}^{d}\frac{2\pi}{n_i}\left|\frac{\partial f}{\partial x_i}\right|+\sum_{1\leqslant i\leqslant j\leqslant d}\frac{2\pi}{n_i}\frac{2\pi}{n_j} \\
&\qquad\times\left|\frac{\partial^2 f}{\partial x_i x_j}\right|+\cdots+\frac{(2\pi)^d}{n_1\cdots n_d}\left|\frac{\partial^d f}{\partial x_1\cdots\partial x_d}\right|\Bigg\}\,dx_1\cdots dx_d. \qquad (10)
\end{aligned}$$

For a univariate function, $g(x_1)$, Hölder's inequality with index $1/p_1+1/q_1=1$ implies that

$$\begin{aligned}
&\left(\sum_{k_1=1}^{n_1}\left(\int_{v_{1,k_1-1}}^{v_{1,k_1}}|g(x_1)|\,dx_1\right)^{p_1}\right)^{1/p_1} \\
&\quad\leqslant\left(\sum_{k_1=1}^{n_1}\left(\frac{2\pi}{n_1}\right)^{p_1/q_1}\int_{v_{1,k_1-1}}^{v_{1,k_1}}|g(x_1)|^{p_1}\,dx_1\right)^{1/p_1}\leqslant\left(\frac{2\pi}{n_1}\right)^{1/q_1}\|g\|_{p_1}. \qquad (11)
\end{aligned}$$

For higher dimensions the same argument gives

$$\begin{aligned}
&\left(\sum_{k_d=1}^{n_2}\left(\cdots\left(\sum_{k_1=1}^{n_1}\left(\int_{v_{1,k_1-1}}^{v_{1,k_1}}\cdots\int_{v_{d,k_d-1}}^{v_{d,k_d}}|g|\,dx_1\cdots dx_d\right)^{p_1}\right)^{p_2/p_1}\cdots\right)^{p_d/p_{d-1}}\right)^{1/p_d} \\
&\quad\leqslant\left(\frac{2\pi}{n_1}\right)^{1/q_1}\cdots\left(\frac{2\pi}{n_1}\right)^{1/q_d}\|g\|_{\mathbf{p}}. \qquad (12)
\end{aligned}$$

For the case $d=1$, applying the Minkowski inequality to (9) and then applying (11) with $g=f$ and $g=f'$ gives

$$\begin{aligned}
\left(\sum_{k_1=1}^{n_1}\left|f(v_{1,k_1})\right|\right)^{1/p_1} &\leqslant \frac{n_1}{2\pi}\left(\frac{2\pi}{n_1}\right)^{1/q_1}\left(\sum_{k_1=1}^{n_1}\left(\int_{v_{1,k_1-1}}^{v_{1,k_1}}|f(x_1)|\,dx_1\right)^{p_1}\right)^{1/p_1} \\
&\quad+\left(\frac{2\pi}{n_1}\right)^{1/q_1}\left(\sum_{k_1=1}^{n_1}\left(\int_{v_{1,k_1-1}}^{v_{1,k_1}}|f'(x_1)|\,dx_1\right)^{p_1}\right)^{1/p_1} \\
&=\left(\frac{2\pi}{n_1}\right)^{-1/p_1}\|f\|_{p_1}+\left(\frac{2\pi}{n_1}\right)^{1/q_1}\|f'\|_{p_1}.
\end{aligned}$$

The same argument for arbitrary d applied to (10) and (12) completes the proof of this lemma. □

Lemma 7 (*Nikolskii [18, p. 227, Theorem 5.6.3]*). *Let* $\boldsymbol{\ell}=(\ell_1,\ldots,\ell_d)\in\mathbb{Z}_+^d$, $\mathbf{r}=(r_1,\ldots,r_d)\in\mathbb{R}_+^d$, $\mathbf{1}\leqslant\mathbf{p}\leqslant\infty$, $1-\sum_{k=1}^d\ell_k/r_k>0$, and $\mathbf{r}':=(1-\sum_{k=1}^d\ell_k/r_k)\mathbf{r}$. For $f\in B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)$ *it follows that there exists a constant C depending on* $\mathbf{r},\mathbf{r}',\boldsymbol{\ell},\mathbf{p}$, *and* θ, *but independent of f, such that*

$$\|f^{(\boldsymbol{\ell})}\|_{B^{\mathbf{r}'}_{\mathbf{p}\theta}(\mathbb{T}^d)} \leqslant C\|f\|_{B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)}.$$

Lemma 8. *Let* $f \in L^{\mathbf{1}}_{\mathbf{p}}(\mathbb{T}^d)$, $\{f(\mathbf{k}\pi/\mathbf{N})\} \in l_{\mathbf{p}}$, $\mathbf{1} < \mathbf{p} < \infty$. *Then for every trigonometric polynomial* $t \in T(\mathbf{N}, d)$ *the error for Dirichlet interpolation is bounded by*

$$\|f - \mathcal{D}_{\mathbf{N}}(f)\|_{\mathbf{p}} \leqslant C_{\mathbf{p}}\|(f-t)(\mathbf{2k}\pi/(\mathbf{2N}+\mathbf{1}))\|_{\mathbf{p},\mathbf{2N}} + \|f-t\|_{\mathbf{p}}.$$

Proof. Let $t \in T(\mathbf{N}, d)$, in view of relation (5), $\mathcal{D}_{\mathbf{N}}(t)(\mathbf{x}) \equiv t(\mathbf{x})$. Then it follows that

$$\begin{aligned} |f - \mathcal{D}_{\mathbf{N}}(f)\|_{\mathbf{p}} &\leqslant \|\mathcal{D}_{\mathbf{N}}(f) - \mathcal{D}_{\mathbf{N}}(t)\|_{\mathbf{p}} + \|f-t\|_{\mathbf{p}} = \|\mathcal{D}_{\mathbf{N}}(f-t)\|_{\mathbf{p}} + \|f-t\|_{\mathbf{p}} \\ &\leqslant C_{\mathbf{p}}\|(f-t)(\mathbf{2k}\pi/(\mathbf{2N}+\mathbf{1}))\|_{\mathbf{p},\mathbf{2N}} + \|f-t\|_{\mathbf{p}}. \qquad \square \end{aligned}$$

Lemma 9. *Suppose that* $f \in B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)$ *and* $\mathbf{N} = (\mathbf{k}+\mathbf{2})(\mathbf{n}-\mathbf{1})$, $\mathbf{1} < \mathbf{p} < \infty$, *where the equation* $\mathbf{N} = (\mathbf{k}+\mathbf{2})(\mathbf{n}-\mathbf{1})$ *means that* $\mathbf{N} = (N_1, \ldots, N_d)$, $\mathbf{n} = (n_1, \ldots, n_d)$, $N_j = (k_j+2)(n_j-1)$, *for all* $j = 1, \ldots, d$, *where* n_j *is defined as the integer part of* $n^{g(\mathbf{r})/r_j}$. *Then it follows that*

$$\|f - \mathcal{D}_{\mathbf{N}}(f)\|_{\mathbf{p}} \ll n^{-g(\mathbf{r})}\|f\|_{B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)}$$

and

$$\|f - \mathcal{D}_{\mathbf{N}}(f)\|_{L^{\mathbf{s}}_{\mathbf{p}}(\mathbb{T}^d)} \ll n^{-g(\mathbf{r})(1-\beta)}\|f\|_{B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)}.$$

Proof. First consider the case $d = 2$. The argument in higher dimensions is analogous. Let

$$T_{r,n}(u) = \left(\frac{\sin\frac{nu}{2}}{n\sin\frac{u}{2}}\right)^{2r+4} \quad \text{and} \quad \gamma_{r,n} = \int_{-\pi}^{\pi} T_{r,n}(u)\,du.$$

The function $T_{r,n}(u)$ is a trigonometric polynomial of the degree $(r+2)(n-1)$, and

$$\frac{1}{\gamma_{r,n}}\int_0^{\pi} T_{r,n}(u)u^k\,du \leqslant C(k)n^{-k}, \quad k = 0, 1, \ldots, 2r+3 \tag{13}$$

(see [29, Chapter 1, Section 1]). Let

$$\begin{aligned} \sigma_{(k_1,n_1)(k_2,n_2)}(f; x_1, x_2) := \frac{1}{\gamma_{k_1,n_1}\gamma_{k_2,n_2}} &\int_0^{2\pi}\int_0^{2\pi} T_{k_1,n_1}(t_1)T_{k_2,n_2}(t_2) \\ &\times \sum_{i=1}^{k_1}\sum_{j=1}^{k_2}(-1)^{i+j}\binom{k_1}{i}\binom{k_2}{j} \\ &\times f(x_1+it_1, x_2+jt_2)\,dt_1\,dt_2. \end{aligned}$$

Then one may express $\sigma_{(k_1,n_1)(k_2,n_2)}(f; x_1, x_2) - f(x_1, x_2)$ as the sum of $\phi_1(x_1, x_2)$ and $\phi_2(x_1, x_2)$, where

$$\begin{aligned}
\phi_1(x_1, x_2) &= \frac{1}{\gamma_{k_2,n_2}} \int_0^{2\pi} T_{k_2,n_2}(t_2)\Phi_1(x_1, x_2, t_2)\, dt_2, \\
\Phi_1(x_1, x_2, t_2) &= \frac{1}{\gamma_{k_1,n_1}} \int_0^{2\pi} T_{k_1,n_1}(t_1) \\
&\quad \times \left\{ \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} (-1)^{i+j} \binom{k_1}{i} \binom{k_2}{j} f(x_1 + it_1, x_2 + jt_2) \right. \\
&\quad \left. - \sum_{j=1}^{k_2} (-1)^{j-1} \binom{k_2}{j} f(x_1, x_2 + jt_2) \right\} dt_1, \\
\phi_2(x_1, x_2) &= \frac{1}{\gamma_{k_2,n_2}} \int_0^{2\pi} T_{k_2,n_2}(t_2) \\
&\quad \times \left\{ \sum_{j=1}^{k_2} (-1)^{j-1} \binom{k_2}{j} f(x_1, x_2 + jt_2) - f(x_1, x_2) \right\} dt_2.
\end{aligned}$$

The norms of the functions above may be bounded by the generalized Minkowskii inequality [29, p. 10]:

$$\left\| \int_a^b f(\cdot, y)\, dy \right\|_p \leqslant \int_a^b \| f(\cdot, y)\|_p\, dy, \quad 1 \leqslant p \leqslant \infty.$$

Specifically, it follows that

$$\begin{aligned}
\|\phi_1\|_{\mathbf{p}} &= \frac{1}{\gamma_{k_2,n_2}} \left\| \int_0^{2\pi} T_{k_2,n_2}(t_2)\Phi_1(\cdot, \cdot, t_2)\, dt_2 \right\|_{\mathbf{p}} \\
&\leqslant \frac{1}{\gamma_{k_2,n_2}} \int_0^{2\pi} T_{k_2,n_2}(t_2) \|\Phi_1(\cdot, \cdot, t_2)\|_{\mathbf{p}}\, dt_2
\end{aligned}$$

and

$$\begin{aligned}
&\|\Phi_1(\cdot, \cdot, t_2)\|_{\mathbf{p}} \\
&\quad = \frac{1}{\gamma_{k_1,n_1}} \left\| \int_0^{2\pi} T_{k_1,n_1}(t_1) \left\{ \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} (-1)^{i+j} \binom{k_1}{i} \binom{k_2}{j} f(\cdot + it_1, \cdot + jt_2) \right. \right. \\
&\qquad \left. \left. - \sum_{j=1}^{k_2} (-1)^{j-1} \binom{k_2}{j} f(\cdot, \cdot + jt_2) \right\} dt_1 \right\|_{\mathbf{p}} \\
&\quad \leqslant \frac{2^{k_2+1}}{\gamma_{k_1,n_1}} \int_0^{2\pi} T_{k_1,n_1}(t_1)\omega_{k_1}(f; t_1)_{\mathbf{p}}\, dt_1.
\end{aligned}$$

Hence,

$$\|\phi_1\|_{\mathbf{p}} \leqslant \frac{2^{k_2+1}}{\gamma_{k_1,n_1}} \int_0^{2\pi} T_{k_1,n_1}(t_1)\omega_{k_1}(f;t_1)_{\mathbf{p}}\, dt_1.$$

In the same way, it can be shown that

$$\|\phi_2\|_{\mathbf{p}} \leqslant \frac{2}{\gamma_{k_2,n_2}} \int_0^{2\pi} T_{k_2,n_2}(t_2)\omega_{k_2}(f;t_2)_{\mathbf{p}}\, dt_2.$$

Therefore, it follows that

$$\begin{aligned} &\left\|\sigma_{(k_1,n_1)(k_2,n_2)}(f;\cdot,\cdot) - f(\cdot,\cdot)\right\|_{\mathbf{p}} \\ &\quad\leqslant \frac{2^{k_2+1}}{\gamma_{k_1,n_1}} \int_0^{2\pi} T_{k_1,n_1}(t_1)\omega_{k_1}(f;t_1)_{\mathbf{p}}\, dt_1 \\ &\qquad + \frac{2}{\gamma_{k_2,n_2}} \int_0^{2\pi} T_{k_2,n_2}(t_2)\omega_{k_2}(f;t_2)_{\mathbf{p}}\, dt_2. \end{aligned} \tag{14}$$

The first term in this inequality may be bounded using Definition 1 and inequality (13):

$$\begin{aligned} &\int_0^{2\pi} T_{k_1,n_1}(t_1)\omega_{k_1}(f;t_1)_{\mathbf{p}}\, dt_1 \\ &\quad\leqslant \left(\int_0^{2\pi} \left(\frac{\omega_{k_1}(f;t_1)_{\mathbf{p}}}{|t_1|^{r_1+1/\theta}}\right)^{\theta} dt_1\right)^{1/\theta} \\ &\qquad\times \left(\int_0^{2\pi} |t_1|^{(r_1+1/\theta)\theta'} \left|\left(\frac{\sin(n_1t_1/2)}{n_1\sin(t_1/2)}\right)^{2r_1+4}\right|^{\theta'} dt_1\right)^{1/\theta'} \\ &\quad\leqslant c_1 n_1^{-r_1}\, \|f\|_{b^{r_1}_{x_1\mathbf{p}\theta}(\mathbb{T}^2)}\,. \end{aligned}$$

The second term in (14) may also be bounded in the same way. Together with the above inequality this gives

$$\begin{aligned} &\left\|\sigma_{(k_1,n_1)(k_2,n_2)}(f;\cdot,\cdot) - f(\cdot,\cdot)\right\|_{\mathbf{p}} \\ &\quad\leqslant c_1 \left(n_1^{-r_1}\, \|f\|_{b^{r_1}_{x_1\mathbf{p}\theta}(\mathbb{T}^2)} + n_2^{-r_2}\, \|f\|_{b^{r_2}_{x_2\mathbf{p}\theta}(\mathbb{T}^2)}\right). \end{aligned}$$

A similar argument leads to upper bounds on a similar quantity, but involving the partial derivatives of *f*:

$$\begin{aligned} &\left\|\frac{\partial}{\partial x_i}\sigma_{(k_1,n_1)(k_2,n_2)}(f;x_1,x_2) - \frac{\partial}{\partial x_i} f(x_1,x_2)\right\|_{\mathbf{p}} \\ &\quad= \left\|\sigma_{(k_1,n_1)(k_2,n_2)}\left(\frac{\partial f}{\partial x_i}\right) - \frac{\partial f}{\partial x_i}\right\|_{\mathbf{p}} \\ &\quad\leqslant c_1 \left(n_1^{-r_1(1-1/r_i)} \left\|\frac{\partial f}{\partial x_i}\right\|_{b^{r_1(1-1/r_i)}_{x_1\mathbf{p}\theta}(\mathbb{T}^2)} + n_2^{-r_2(1-1/r_i)} \left\|\frac{\partial f}{\partial x_i}\right\|_{b^{r_2(1-1/r_i)}_{x_2\mathbf{p}\theta}(\mathbb{T}^2)}\right), \\ &\qquad i = 1, 2 \end{aligned}$$

and

$$\left\|\frac{\partial^2}{\partial x_1 x_2}\sigma_{(k_1,n_1)(k_2,n_2)}(f; x_1, x_2) - \frac{\partial^2}{\partial x_1 x_2} f(x_1, x_2)\right\|_{\mathbf{p}}$$
$$\leqslant c_1 \left(n_1^{-r_1(1-1/g(r_1,r_2))} \left\|\frac{\partial^2 f}{\partial x_1 x_2}\right\|_{b_{x_1\mathbf{p}\theta}^{r_1(1-1/g(r_1,r_2))}(\mathbb{T}^2)} + \right.$$
$$\left. + n_2^{-r_2(1-1/g(r_1,r_2))} \left\|\frac{\partial^2 f}{\partial x_1 x_2}\right\|_{b_{x_2\mathbf{p}\theta}^{r_2(1-1/g(r_1,r_2))}(\mathbb{T}^2)} \right),$$

where $g(r_1, r_2)$ is the harmonic average of r_1 and r_2 defined in (3).

For arbitrary dimension, d, the arguments above may be generalized to give

$$\|\sigma_{(\mathbf{k},\mathbf{n})}(f) - f\|_{\mathbf{p}} \ll \sum_{m=1}^{d} n_m^{-r_m} \|f\|_{b_{x_m\mathbf{p}\theta}^{r_m}(\mathbb{T}^d)}$$

and

$$\left\|\frac{\partial^{l+s_i}\sigma_{\mathbf{k},\mathbf{n}}(f)}{\partial x_{j_1}\cdots\partial x_{j_l}\partial x_i^{s_i}} - \frac{\partial^{l+s_i} f}{\partial x_{j_1}\cdots\partial x_{j_l}\partial x_i^{s_i}}\right\|_{\mathbf{p}}$$
$$\ll \sum_{m=1}^{d} n_m^{-r_m(1-1/g(r_{j_1},\ldots,r_{j_l})-s_i/r_i)} \|f\|_{b_{x_m\mathbf{p}\theta}^{r_m(1-1/g(r_{j_1},\ldots,r_{j_l})-s_i/r_i)}(\mathbb{T}^d)},$$

where $(\mathbf{k}, \mathbf{n}) = (k_1, n_1), (k_2, n_2), \ldots, (k_d, n_d)$.

The final step in this proof is to choose $t(\mathbf{x}) = \sigma_{(\mathbf{k},\mathbf{n})}(f; \mathbf{x})$ and apply Lemmas 8 and then 6. It follows that

$$\|f - \mathcal{D}_{\mathbf{N}}(f)\|_{\mathbf{p}}$$
$$\leqslant C_{d,\mathbf{p}} \|(f-t)(2\mathbf{k}\pi/(2\mathbf{N}+\mathbf{1}))\|_{\mathbf{p},2\mathbf{N}} + \|f-t\|_{\mathbf{p}}$$
$$\leqslant C_{d,\mathbf{p}} \left(\|f-t\|_{\mathbf{p}} + \sum_{1\leqslant i\leqslant d} \frac{2\pi}{n_i} \left\|\frac{\partial(f-t)}{\partial x_i}\right\|_{\mathbf{p}} + \cdots + \prod_{i=1}^{d} \frac{2\pi}{n_i} \left\|\frac{\partial^d (f-t)}{\partial x_1\cdots\partial x_d}\right\|_{\mathbf{p}} \right)$$
$$\leqslant C_{d,\mathbf{p}} \left(\sum_{m=1}^{d} n_m^{-r_m} \|f\|_{b_{x_i\mathbf{p}\theta}^{r_m}(\mathbb{T}^d)} + \sum_{i=1}^{d} \frac{2\pi}{n_i} \sum_{m=1}^{d} n_m^{-r_m(1-1/g(r_i))} \left\|\frac{\partial f}{\partial x_i}\right\|_{b_{x_m\mathbf{p}\theta}^{r_m(1-1/g(r_i))}(\mathbb{T}^d)} \right.$$
$$\left. + \cdots + \sum_{m=1}^{d} \prod_{i=1}^{d} \frac{2\pi}{n_i} n_m^{-r_m(1-1/g(\mathbf{r}))} \left\|\frac{\partial^d f}{\partial x_1\cdots\partial x_d}\right\|_{b_{x_m\mathbf{p}\theta}^{r_m(1-1/g(\mathbf{r}))}(\mathbb{T}^d)} \right).$$

By virtue of Lemma 7, the quantity above can be further estimated as follows:

$$\|f-\mathcal{D}_{\mathbf{N}}(f)\|_{\mathbf{p}} \leqslant C_{d,\mathbf{p}}\left(\sum_{m=1}^{d} n_m^{-r_m} + \sum_{i=1}^{d}\frac{2\pi}{n_i}\sum_{m=1}^{d} n_m^{-r_m(1-1/g(r_i))} + \sum_{1\leqslant i<j\leqslant d}\frac{2\pi}{n_i}\frac{2\pi}{n_j}\sum_{m=1}^{d} n_m^{-r_m(1-1/g(r_i,r_j))} + \cdots + \sum_{m=1}^{d}\prod_{i=1}^{d}\frac{2\pi}{n_i}n_m^{-r_m(1-1/g(\mathbf{r}))}\right)\|f\|_{B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)}$$

and

$$\left\|\frac{\partial^{s_i} f}{\partial x_i^{s_i}} - \mathcal{D}_{\mathbf{N}}\left(\frac{\partial^{s_i} f}{\partial x_i^{s_i}}\right)\right\|_{\mathbf{p}} \ll \left(\sum_{m=1}^{d} n_m^{-r_m(1-s_i/r_i)} + \sum_{j=1}^{d}\frac{2\pi}{n_j}\sum_{m=1}^{d} n_m^{-r_m(1-1/g(r_j)-s_i/r_i)} + \cdots + \sum_{m=1}^{d}\prod_{j=1}^{d}\frac{2\pi}{n_j}n_m^{-r_m(1-1/g(\mathbf{r})-s_i/r_i)}\right)\|f\|_{B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)}.$$

Let n_j be defined as in the hypothesis of this lemma. The conclusion of the proof now follows. □

Using the methods described in [18, Chapter 6, Section 3; 29, Chapter 2, Section 3; 13] leads to the following lemma:

Lemma 10. *For* $\mathbf{1}\leqslant\mathbf{p}<\mathbf{q}\leqslant\boldsymbol{\infty}$, $1\leqslant\theta\leqslant\infty$, $\gamma = 1-\sum_{i=1}^{d}(1/p_i-1/q_i)1/r_i > 0$, *and* $\mathbf{r}'=\gamma\mathbf{r}$ *it follows that*

$$B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d) \hookrightarrow B^{\mathbf{r}'}_{\mathbf{q}\theta}(\mathbb{T}^d),$$

which means that under the given conditions, if a function f belongs to the space $B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{R}^d)$, *then it also belongs to the space* $f\in B^{\mathbf{r}'}_{\mathbf{q}\theta}(\mathbb{R}^d)$, *and the inequality*

$$\|f\|_{B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)} \leqslant C\|f\|_{B^{\mathbf{r}'}_{\mathbf{p}\theta}(\mathbb{T}^d)}$$

holds, where C is a constant depending only on $\mathbf{p},\mathbf{q},\mathbf{r},\mathbf{r}'$.

Proof of upper bounds in Theorems 2 and 3. From Lemma 9 it follows that if $1<\mathbf{q}\leqslant\mathbf{p}<\boldsymbol{\infty}$, or $\mathbf{q}=\mathbf{1}$, $\mathbf{1}<\mathbf{p}<\boldsymbol{\infty}$, then

$$e_n^L\left(S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d), L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)\right) \leqslant \sup_{f\in S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)} \|f-D_{\mathbf{N}}(f)\|_{L^{\mathbf{s}}_{\mathbf{p}}(\mathbb{T}^d)} \ll n^{-g(\mathbf{r})(1-\beta)}.$$

Furthermore, by virtue of Lemmas 9 and 10, if $1 < \mathbf{p} \leqslant \mathbf{q} < \infty$, or $\mathbf{p} = \mathbf{1}$, $\mathbf{1} < q < \infty$, then

$$\begin{aligned} e_n^L\left(S_{\mathbf{p}\theta}^{\mathbf{r}}(\mathbb{T}^d), L_{\mathbf{q}}^{\mathbf{s}}(\mathbb{T}^d)\right) &\leqslant \sup_{f \in S_{\mathbf{p}\theta}^{\mathbf{r}}(\mathbb{T}^d)} \|f - D_{\mathbf{N}}(f)\|_{L_{\mathbf{q}}^{\mathbf{s}}(\mathbb{T}^d)} \\ &\ll \sup_{f \in S_{\mathbf{q}\theta}^{\mathbf{r}'}(\mathbb{T}^d)} \|f - D_{\mathbf{N}}(f)\|_{L_{\mathbf{q}}^{\mathbf{s}}(\mathbb{T}^d)} \\ &\ll n^{-g(\mathbf{r})(1-\beta)+g(\mathbf{r})\sum_{i=1}^d (1/p_i - 1/q_i)1/r_i}, \end{aligned}$$

which gives the estimates of upper bound of Theorems 2 and 3. $\square$

4. Lower bounds

The lower bounds in Theorems 2 and 3 are obtained by constructing suitable bump functions. This approach is described in [5] and elsewhere.

Lemma 11. *Let* $\mathbf{k} = (k_1, k_2, \ldots, k_d) \in \mathbb{N}^d$, $\mathbf{r} = (r_1, r_2, \ldots, r_d) \in \mathbb{R}_+^d$, $\mathbf{s} = (s_1, s_2, \ldots, s_d) \in \mathbb{N}^d$, $k_i > r_i > s_i$, $i = 1, \ldots, d$, $1 \leqslant \mathbf{p} \leqslant \infty$, $1 \leqslant \theta \leqslant \infty$. *Then for any points* $B_n = \{\mathbf{b}_1 \cdots \mathbf{b}_n\} \subset \mathbb{T}^d$, *there exists a function* $f \in C^\infty(\mathbb{T}^d) \bigcap B_{\mathbf{p}\theta}^{\mathbf{r}}(\mathbb{T}^d)$ *with support in* $\mathbb{T}_0$, $\mathbb{T}_0$ *is the interior of* $\mathbb{T}$, *such that*

$$f(\mathbf{b}_j) = 0, \quad j = 1, \ldots, n$$

and satisfies the following conditions:

$$\|f\|_{L_{\mathbf{q}}^{\mathbf{s}}(\mathbb{T}^d)} \gg n^{-g(\mathbf{r})(1-\beta)}, \quad 1 \leqslant \mathbf{q} \leqslant \mathbf{p} \leqslant \infty; \tag{15a}$$

$$\|f\|_{L_{\mathbf{q}}^{\mathbf{s}}(\mathbb{T}^d)} \gg n^{-g(\mathbf{r})(1-\beta)+g(\mathbf{r})\sum_{i=1}^d (1/p_i - 1/q_i)1/r_i}, \quad 1 \leqslant \mathbf{p} \leqslant \mathbf{q} \leqslant \infty. \tag{15b}$$

Proof. For any given $n \in \mathbb{N}$ let $m_i := \lfloor n^{g(\mathbf{r})/r_i} \rfloor$, $i = 1, \ldots, d$, so that $\widetilde{n} := \prod_{i=1}^d m_i \asymp n$. The torus $\mathbb{T}^d$ is subdivided into $\overline{n} := 2^d \widetilde{n}$ equal closed rectangles $\{Q_j\}$ with common side length vector $h = (\pi m_1^{-1}, \ldots, \pi m_d^{-1})$ and mutually disjoint interiors:

$$Q_j := \{\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{T}^d : c_i^{(j)} \leqslant x_i \leqslant c_i^{(j)} + m_i^{-1}\pi,\ i = 1, \ldots, d\}.$$

Define a single bump $\phi(t)$ to be a fixed function in $C^\infty(\mathbb{R})$ with support contained in $\mathbb{T}_0$, such that $0 \leqslant \phi(t) \leqslant 1$, $t \in \mathbb{T}$, $\phi(t) = 1$, when $t \in [\pi/2, 3\pi/2]$, and $\|\phi^{(s_i)}\|_{p_i} \geqslant c > 0$, $i = 1, \ldots, d$, where the positive constant c depends on $\mathbf{p}$, $\mathbf{s}$. The function f_j is defined to have a bump only in the rectangle Q_j as follows:

$$f_j(\mathbf{x}) = \prod_{i=1}^d m_i^{1/p_i} (2n)^{-g(\mathbf{r})} \phi(2m_i(x_i - c_i^{(j)})), \quad j = 1, \ldots, \overline{n}.$$

For any $\mathbf{p}$ these functions have the same norm:

$$\|f_j\|_{\mathbf{p}} = (2n)^{-g(\mathbf{r})} \prod_{i=1}^{d} \|\phi\|_{L_{p_i}(\mathbb{T})}, \quad j = 1, \ldots, \overline{n}$$

and

$$\begin{aligned}\|f_j\|_{L^{\mathbf{s}}_{\mathbf{p}}(\mathbb{T}^d)} &= (2n)^{-g(\mathbf{r})} \left(\prod_{i=1}^{d} \|\phi\|_{L_{p_i}(\mathbb{T})} + \sum_{k=1}^{d} m_k^{s_k} \prod_{i=1}^{d} \|\phi\|_{L_{p_i}(\mathbb{T})} \frac{\|\phi^{(s_k)}\|_{L_{p_k}(\mathbb{T})}}{\|\phi\|_{L_{p_k}(\mathbb{T})}} \right) \\ &\gg n^{-g(\mathbf{r})} \max_{1\leqslant i\leqslant d} \{m_i^{s_i}\}. \end{aligned} \tag{16}$$

For $1 \leqslant i \leqslant d$, the Minkowskii inequality implies that

$$\begin{aligned}&\|\Delta_{h_i}^{k_i} f_j\|_{\mathbf{p}} \\ &\quad = \left\| \int_0^{h_i} \cdots \int_0^{h_i} \frac{\partial^{k_i}}{\partial x_i} f_j(x_1, \ldots, x_i + u_1 + \cdots + u_{k_i}, x_{i+1}, \cdots, x_d)\, du_1 \cdots du_{k_i} \right\|_{\mathbf{p}} \\ &\quad \ll (2n)^{-g(\mathbf{r})} (m_i|h_i|)^{k_i} \prod_{j=1}^{d} \|\phi\|_{L_{p_j}(\mathbb{T})} \frac{\|\phi^{(k_i)}\|_{L_{p_i}(\mathbb{T})}}{\|\phi\|_{L_{p_i}(\mathbb{T})}},\end{aligned}$$

independent of j, which implies that $\omega_{k_i}(f_j, h_i)_{\mathbf{p}} \ll (2n)^{-g(\mathbf{r})}(m_i|h_i|)^{k_i}$. On the other hand, it is also true that $\|\Delta_{h_i}^{k_i} f_j\|_{\mathbf{p}} \ll \|f_j\|_{\mathbf{p}}$, and therefore, $\omega_{k_i}(f_j, t_i)_{\mathbf{p}} \ll \|f_j\|_{\mathbf{p}} \asymp (2n)^{-g(\mathbf{r})}$. Thus, for $1 \leqslant \theta < \infty$ it follows that

$$\begin{aligned}\|f_j\|_{b^{r_i}_{x_i\mathbf{p}\theta}\mathbb{T}^d} &\ll (2n)^{-g(\mathbf{r})} \left(\int_0^{m_i^{-1}} m_i^{k_i\theta} t^{(k_i - r_i)\theta - 1}\, dt + \int_{m_i^{-1}}^{\infty} t^{-r_i\theta - 1}\, dt \right)^{1/\theta} \\ &\ll (2n)^{-g(\mathbf{r})} m_i^{r_i} \ll 1, \quad i = 1, \ldots, d, \ j = 1, \ldots, \overline{n}. \end{aligned} \tag{17}$$

This equation can also be proved for the case $\theta = \infty$ in the same way. From inequality (17), it follows that there exists some positive constant c_0 such that $c_0 f_j \in S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)$ for all j.

Now these bump functions are combined together. For any set of n points in the d-dimensional torus, $B_n = \{\mathbf{b}_1, \ldots \mathbf{b}_n\} \subset \mathbb{T}^d$, define J as the set of j for which the interior of Q_j contains none of the points in B_n, i.e., $J = \{j : 1 \leqslant j \leqslant \overline{n},\ B_n \cap Q_{j,0} = \emptyset\}$. Clearly, the cardinality of J is bounded below by $|J| \geqslant \overline{n} - n = (2^d - 1)\widetilde{n} \asymp n$. Since $|J| > \widetilde{n} = \prod_{i=1}^{d} m_i$, one may choose $\widetilde{n}$ of the f_j with $j \in J$ and relabel them as $f_{i,l}$, $l = 1, \ldots, m_i$, $i = 1, \ldots, d$. For any $\boldsymbol{\beta} = (\beta_{1,1}, \ldots, \beta_{1,m_1}, \ldots, \beta_{d,1}, \ldots, \beta_{d,m_d}) \in \mathbb{R}^{\widetilde{n}}$ one may define the following linear combination of bump functions:

$$f_{\boldsymbol{\beta}}(\mathbf{x}) := \prod_{i=1}^{d} \sum_{l=1}^{m_i} \beta_{i,l} c_0 f_{i,l}(\mathbf{x}).$$

From this definition it is easy to see that $\|f_{\boldsymbol{\beta}}\|_{\mathbf{p}} = \|\boldsymbol{\beta}\|_{\mathbf{p}}\|c_0 f_{1,1}\|_{\mathbf{p}}$ and likewise $\|f_{\boldsymbol{\beta}}\|_{L^{\mathbf{s}}_{\mathbf{p}}(\mathbb{T}^d)} = \|\boldsymbol{\beta}\|_{\mathbf{p}}\|c_0 f_{1,1}\|_{L^{\mathbf{s}}_{\mathbf{p}}(\mathbb{T}^d)}$, where

$$\|\boldsymbol{\beta}\|_{\mathbf{p}} = \prod_{i=1}^{d}\left(\sum_{l=1}^{m_i}|\beta_{i,l}|^{p_i}\right)^{1/p_i}.$$

Therefore, for each function $f_{\boldsymbol{\beta}}$, $\boldsymbol{\beta} \in \mathbb{R}^{\widetilde{n}}$ with $\|\boldsymbol{\beta}\|_{\mathbf{p}} \leqslant 1$ belongs to $S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)$. By the Hölder inequality, there is $\overline{\boldsymbol{\beta}} \in \mathbb{R}^{\widetilde{n}}$ such that $\|\overline{\boldsymbol{\beta}}\|_{\mathbf{p}} = 1$.

For any $\mathbf{1} \leqslant \mathbf{q}, \mathbf{p} \leqslant \infty$ it can be shown that

$$\|\overline{\boldsymbol{\beta}}\|_{\mathbf{q}} = \max_{\|\boldsymbol{\beta}\|_{\mathbf{p}} \leqslant 1}\|\boldsymbol{\beta}\|_{\mathbf{q}} = \prod_{i=1}^{d} m_i^{(1/q_i - 1/p_i)_+}. \tag{18}$$

First, consider the case of $\mathbf{q} \leqslant \mathbf{p}$. In this case, $\|\overline{\boldsymbol{\beta}}\|_{\mathbf{q}} = 1$ and $B^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d) \hookrightarrow B^{\mathbf{r}}_{\mathbf{q}\theta}(\mathbb{T}^d)$, therefore, it follows from $f_{\overline{\boldsymbol{\beta}}} \in S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d)$ that there exists a constant $\overline{c_0}$ such that $\overline{c_0} f_{\overline{\boldsymbol{\beta}}} \in S^{\mathbf{r}}_{\mathbf{q}\theta}(\mathbb{T}^d)$. To estimate the $L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)$-norm of this function, note by (16) and (18) that

$$\begin{aligned}\|\overline{c_0} f_{\overline{\boldsymbol{\beta}}}\|_{L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)} &\asymp \|f_{\overline{\boldsymbol{\beta}}}\|_{L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)} = \|\overline{\boldsymbol{\beta}}\|_{\mathbf{q}}\|c_0 f_{1,1}\|_{L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)} = \|c_0 f_{1,1}\|_{L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)}\\ &\gg n^{-g(\mathbf{r})}\max_{1\leqslant i\leqslant d}\{m_i^{s_i}\} \asymp n^{-g(\mathbf{r})}\max_{1\leqslant i\leqslant d}\{n^{g(\mathbf{r})s_i/r_i}\} = n^{-g(\mathbf{r})(1-\beta)}.\end{aligned}$$

This completes the proof of (15a).

If $\mathbf{p} \leqslant \mathbf{q}$, then it is sufficient to simply choose one $j_0 \in J$ and define c_0 such that $c_0 f_{j_0} \in S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T})^d$. By Eqs. (16) and (18) it follows that

$$\begin{aligned}&\|f_{j_0}\|_{L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)}\\ &\quad = (2n)^{-g(\mathbf{r})}\left(\prod_{j=1}^{d}\|\phi\|_{L_{q_i}(\mathbb{T})} + \sum_{i=1}^{d} m_i^{s_i+1/p_i-1/q_i}\prod_{j=1}^{d}\|\phi\|_{L_{q_j}(\mathbb{T})}\frac{\|\phi^{(s_i)}\|_{L_{q_i}(\mathbb{T})}}{\|\phi\|_{L_{q_i}(\mathbb{T})}}\right)\\ &\quad \gg n^{-g(\mathbf{r})(1-\beta)+g(\mathbf{r})\sum_{i=1}^{d}(1/p_i-1/q_i)1/r_i},\end{aligned}$$

which completes the proof (15b). $\square$

Proof of lower bounds in Theorems 2 and 3. If $\mathbf{q} \leqslant \mathbf{p}$, it follows from Lemma 11 that for any given design $B_n = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\} \subset \mathbb{T}^d$, there is a function $f_{\overline{\boldsymbol{\beta}}} \in S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T})^d$, depending on the design, that vanishes on the design, i.e., $f_{\overline{\boldsymbol{\beta}}}(\mathbf{b}_j) = 0$ for $j = 1, \ldots, n$, and has $\|f_{\overline{\boldsymbol{\beta}}}\|_{L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)} \gg n^{-g(\mathbf{r})(1-\beta)}$. Therefore the nth minimum information diameter defined in (1) is bounded below by

$$\begin{aligned}\Delta_n(S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d), L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)) &\gg d(I^{-1}_{B_n}(I_{B_n} f_{\overline{\boldsymbol{\beta}}}) \cap S^{\mathbf{r}}_{\mathbf{p}\theta}(\mathbb{T}^d), L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d))\\ &\gg \|f_{\overline{\boldsymbol{\beta}}}\|_{L^{\mathbf{s}}_{\mathbf{q}}(\mathbb{T}^d)} \gg n^{-g(\mathbf{r})(1-\beta)},\end{aligned}$$

which provides the lower bounds in Theorem 2. The lower bounds for Theorem 3 are proved in a similar way. $\square$

Acknowledgments

This research was completed while Gensun Fang was visiting Hong Kong Baptist University. He is grateful for the hospitality shown to him. The authors thank Drs. Fanglun Huang and Guanggui Chen for their valuable comments on this work. The authors are extremely grateful to the associate editor and the referees for their excellent suggestions, which greatly helped us to improve this paper.

References

[1] K.I. Babenko, Foundation of Numerical Analysis, Nauka, Moskow, 1986 (in Russian).
[2] P.G. Ciarlet, The finite element method for elliptic problems, Studia Math. 53 (1975) 277–302.
[3] S. Dahlke, R.A. DeVore, Besov regularity for elliptic boundary value problems, Comm. Partial Differential Equations 22 (1997) 1–16.
[4] Gensun Fang, Whittaker–Kotelnikov–Shannon sampling theorem and aliasing error, J. Approx. Theory 85 (1996) 115–131.
[5] S. Heinrich, Random approximation in numerical analysis, in: K.D. Bierstedt et al., Function Analysis: Proceedings of the Essen Conference, Lecture Notes in Pure and Applied Mathematics, vol. 150, Marcel Dekker, New York, Bessel, Hong Kong, 1994, pp. 123–171.
[6] S. Heinrich, J.D. Kern, Parallel information-based complexity, J. Complexity 7 (1991) 339–370.
[7] V.H. Hristov, On convergence of some interpolation processes in integral discrete norms, Constructive Function Theory, vol. 81, in: Proceedings of the International Conference on Constructive Function Theory, Varna, June 1–5, 1981, Sofia, 1983, pp. 185–188 (in Russian).
[8] K.G. Ivanov, On the rates of convergence of two moduli of functions, Prisca Stud. Math. Bulgar. 15 (1983) 97–104.
[9] A.N. Kolmogorov, Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse, Ann. Math. 37 (1936) 107–110.
[10] S.N. Kudryavtsev, Recovering a function with its derivatives from function values at a given number of points, Russian Aca. Sci. Izv. Math. 45 (1995) 505–528.
[11] S.N. Kudryavtsev, The best accuracy of reconstruction of finitely smooth functions from their values at a given number of points, Izv. RAN. Math. 62 (1998) 19–53.
[12] J.B. Luo, Y.S. Sun, Optimal recovery and widths of an anisotropic Sobolev class of multivariate functions, Adva. Math. 27 (1998) 69–77 (in Chinese).
[13] G.G. Magaril-Il'yaev, Average-widths of Sobolev classes on $\mathbb{R}^n$, J. Approx. Theory 76 (1994) 65–76.
[14] C.A. Micchelli, A. Pinkus, Some problems in the approximation of functions of two variables and n-widths of integral operators, J. Approx. Theory 24 (1978) 51–77.
[15] C.A. Michelli, T.J. Rivlin, A survey of optimal recovery, in: C.A. Michelli, T.J. Rivlin (Eds.), Optimal Estimates in Approximation Theory, Plenum, New York, 1977, pp. 1–54.
[16] C.A. Michelli, T.J. Rivlin, Lecture on optimal recovery, Lecture Notes in Mathematics, vol. 1129, Springer, Berlin, New York, 1985, pp. 21–93.
[17] S.M. Nikolskii, Approximation of functions of real variable by polynomials, in: Mathematics in USSR for Thirty Years, 1917–1947, M.-L., 1948 (in Russian).
[18] S.M. Nikolskii, Approximation of Functions of Several Variables and Imbedding Theorem, Springer, Berlin, Heidelberg, New York, 1975.
[19] E. Novak, Deterministic and stochastic error bound in numerical analysis, Lecture Notes in Mathematics, vol. 1349, Springer, Berlin, 1988.
[20] A. Pinkus, N-widths in Approximation Theory, Springer, New York, 1985.
[21] K. Ritter, Average-case analysis of numerical problems, Lecture Notes in Mathematics, vol. 1733, Springer, Berlin, 2000.
[22] K.V. Runovski, H.-J. Schmeisser, Marcinkiewicz–Zygmund type inequalities for irregular knots and mixed metrics, Vestnik RUDN Math. 4,5 (1997/98) 90–115.

[23] H.-J. Schmeisser, An unconditional basis in periodic spaces with dominating mixed smoothness properties, Anal. Math. 13 (1987) 153–168.
[24] H.-J. Schmeisser, W. Sickel, Sampling theory and function spaces, in: Applied Mathematics, vol. 1, World Scientific, Singapore, 2001, pp. 205–284.
[25] H.-J. Schmeisser, H. Triebel, Topics in Fourier Analysis and Function Spaces, Wiley, Chichester, 1987.
[26] V.N. Temlyakov, Approximation of Functions with Bounded Mixed Derivative, Nauka, Moskow, 1986; (English transl. in Proceedings of Steklov Inst., American Mathematical Society, Providence, RI, 1989).
[27] V.N. Temlyakov, Approximate recovery of functions of several variables, Mat Sbornik 128 (1985) 256–268 (English translation in Math. USSR Sbornik 56(1987) 249–261).
[28] V.N. Temlyakov, On approximate recovery of functions with bounded mixed derivative, J. Complexity 9 (1993) 41–59.
[29] V.N. Temlyakov, Approximation of Periodic Functions, Nova Science, New York, 1993.
[30] J.F. Traub, G.W. Wasilkowski, H. Woźniakowski, Information-based Complexity, Academic Press, New York, 1988.
[31] J.F. Traub, H. Woźniakowski, A General Theory of Optimal Algorithms, Academic Press, New York, 1980.